Twitter Analysis on Bank of America

Chenxin Xie, Clement Bannem, Philipp Borchert

2/6/2020

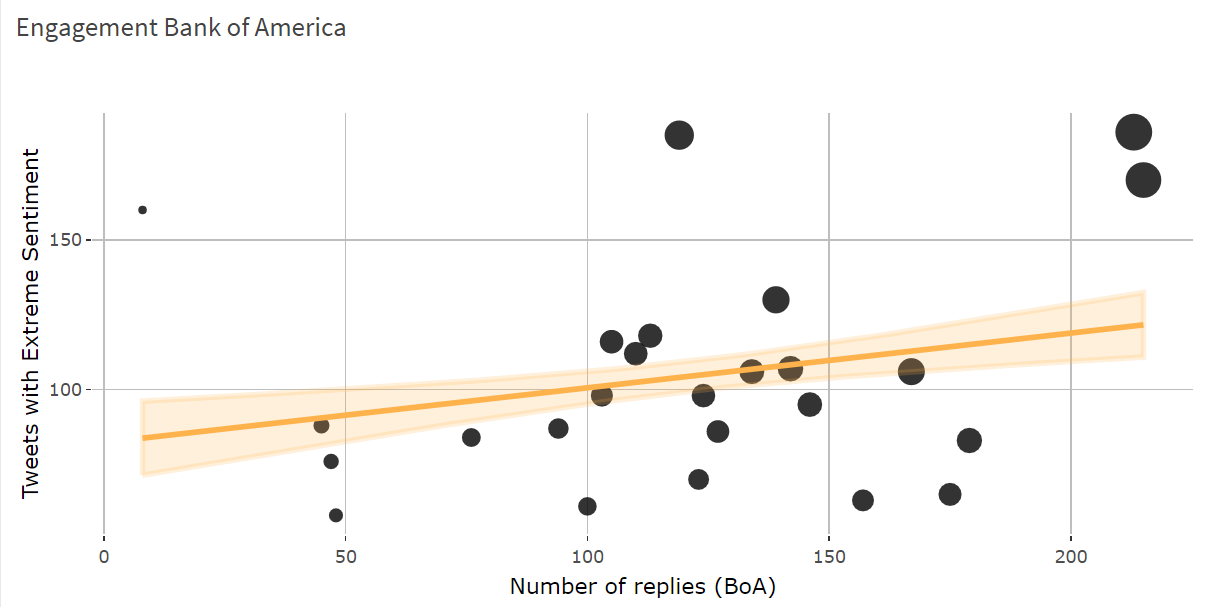
## Project Definition

Our client is Bank of America and this project is aimed to help Bank of America to understand what do people think about their service through analyzing information from one of the most popular social media: Twitter.  
- Twitter API Calling - Data preparation - WordCloud - Sentiment Analysis - Topic Modelling - Follower Profile Summary

## Data Summary

### Tweets posted by Bank of Amercia:

We collected tweets posted by 7 Twitter accounts that represent Bank of America : **@BofA\_News**, **@BofA\_Help**, **@BofA\_Careers**, **@BofA\_Community**, **@BofA\_Tips**, **@BankofAmerica**, **@MerrillLynch**.  
In total, we have collected **16965** tweets from Bank of America Accounts. *Because of the limitation of get\_timeline, we can only collect maximum 3200 tweets per account.*  
With these tweets we have collected, we summarized twitter behavior of Bank of America and some response from people (likes, retweets, etc.). We have found the most common words in the tweets and plotted a wordCloud. We have analyzed Bank of Americas Twitter - Engagement & Activity in regards to the sentiment (positive / negative). According to the subsample we selected, it looks like BoA replies to tweets more often, when there is extreme positive or negative sentiment. (There is no statistical significance due to a lack of datapoints.)



Plot Sentiment/Replies.

### Tweets mentioned about Bank of Amercia:

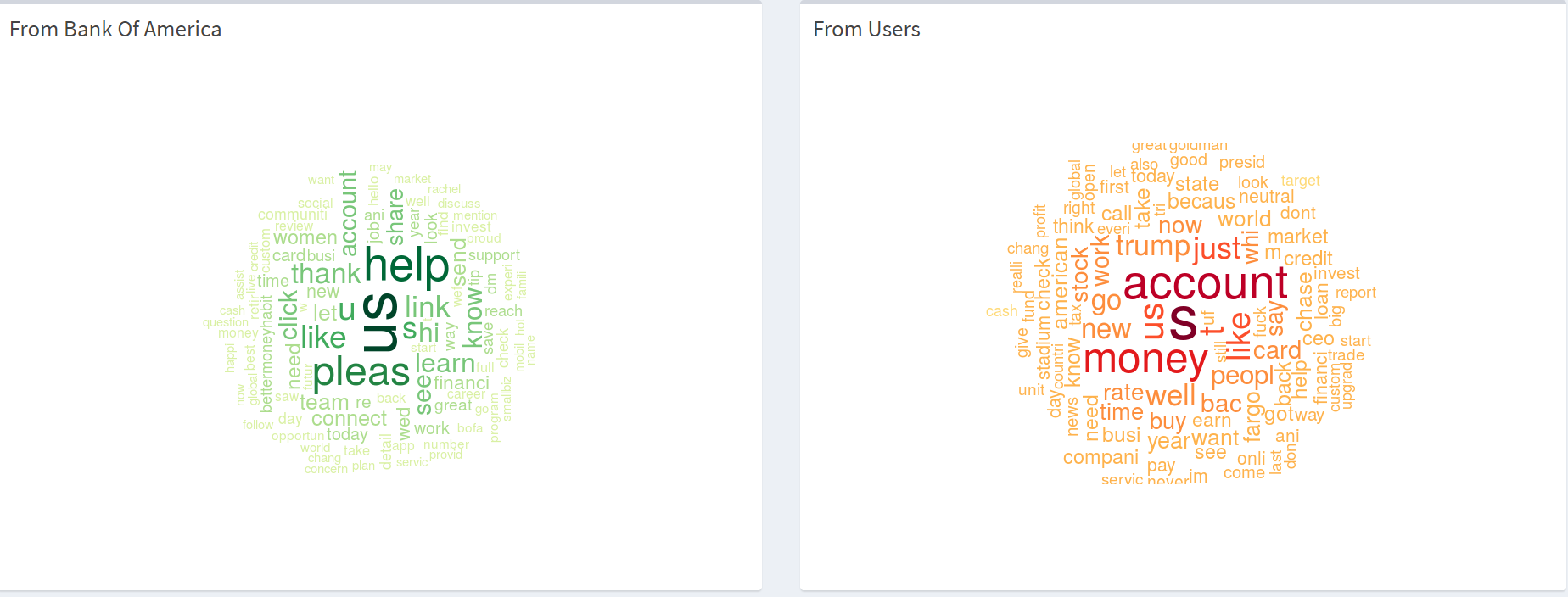
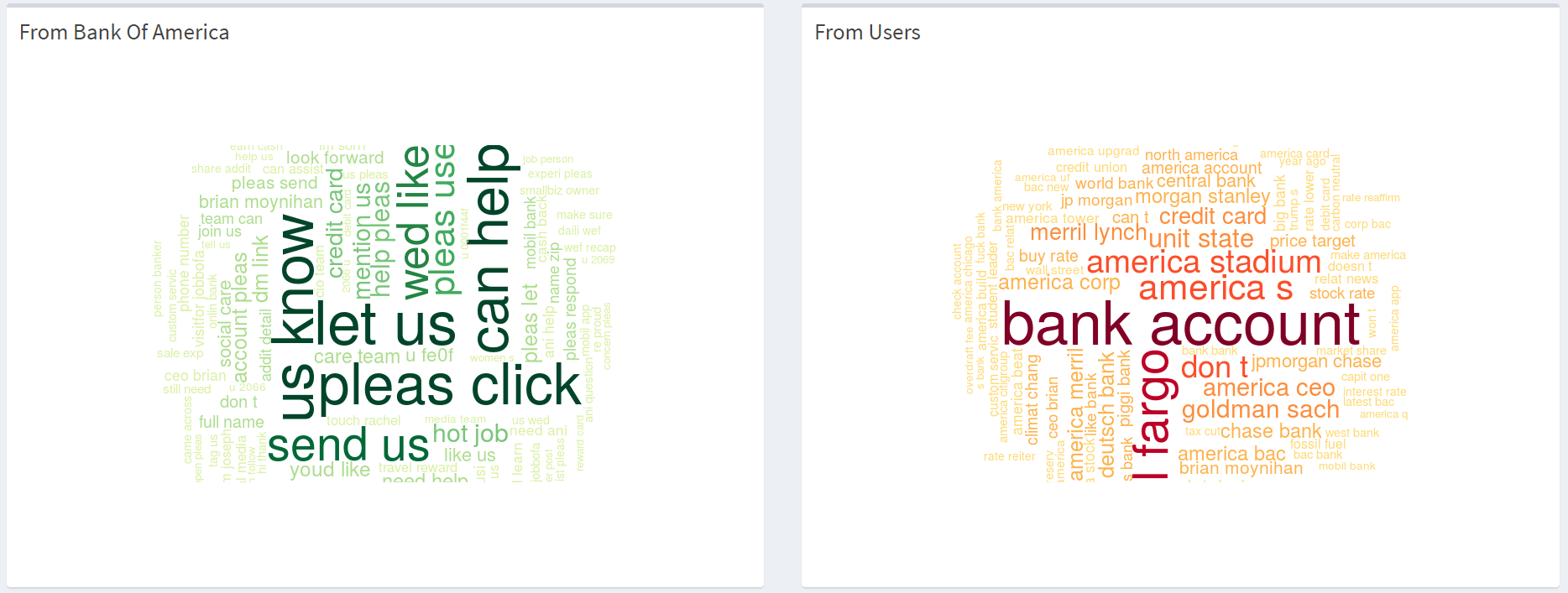
We used **Bank of America**, **#BankofAmerica**, **BofA**, **BAC** as search term to collect the tweets people posted talking about Bank of America.  
We have collected **11676** tweets in total, from **1/1/2020** to **1/31/2020** from Users talking about Bank of America.  
With these tweets we have collected, we plotted the wordCloud, conducted sentiment analysis and topic modelling to discover how do people think about their service and what do people mostly talk about when they mention Bank of America. We have analyzed the language they use and plotted the sentiment over the time period.  
To analyze the data, we used 2 different and 1 combined approach:  
Dictionary lookup: The Sentiment is scored on a scale from -5 to 5 (later rescaled to -1,1) based on the word in the dictionary (afinn).  
Machine Learning: We labelled 1500 tweets from the dataset in regards to their Sentiment. Positive = 1, Neutral = 0, Negative = -1. We trained a Random Forest Model to predict the probability of each tweet belonging to one of the specified Sentiment levels. The score is computed by multiplying the probabilities with their level and adding the numbers. E.g. Predicted (1=0.2, 0=0.1, -1=0.7) becomes 1 \* 0.2 + 0 \* 0.1 + -1 \* 0.7 = -0.5 as a Sentiment Score. Combined Approach: After assigning the dictionary and Machine Learning Sentiment to the tweets, the combined column is computed as the average of the 2 approaches.

### People following Bank of America on Twitter:

**@BankofAmerica** has 540.3K followers.  
We have collected **24424** follower profiles.  
With these profiles we have collected, we summarized the followers’ twitter behavior, plotted location map and ranked the countries of followers with high activity.

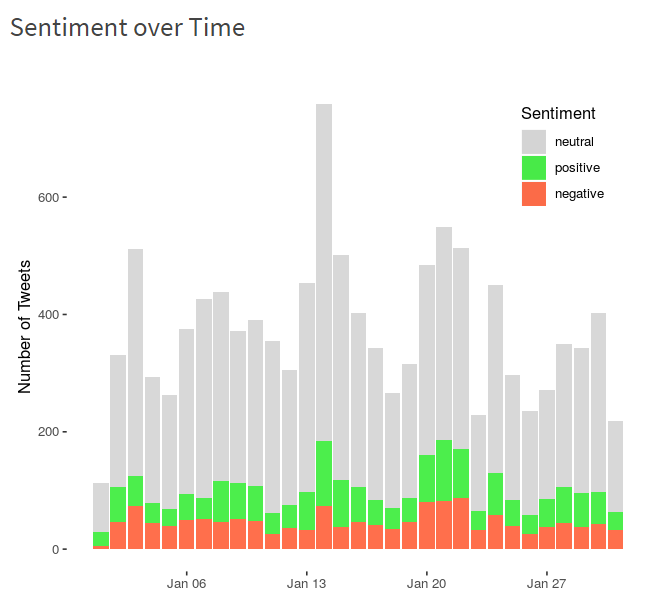
## WordCloud

To show the most common words in the tweets, we made wordClouds based on unigram and bigram for both tweets posted by Bank of America and tweets mention Bank of America.

## Sentiment Analysis

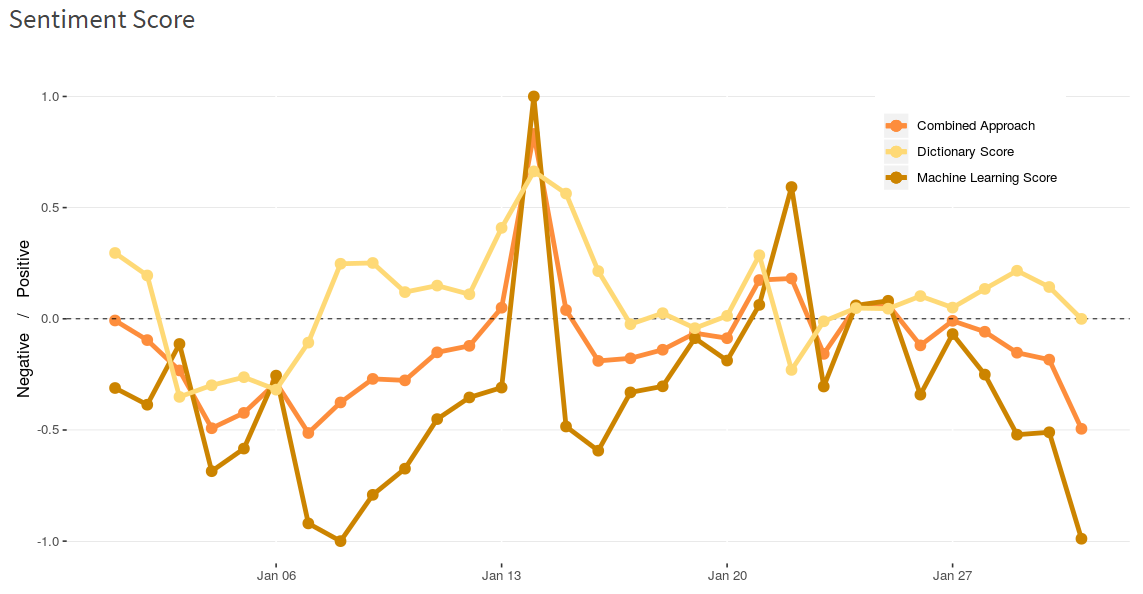
To discover how do users feel about Bank of America, we conducted Sentiment Analysis on the tweets people talk about Bank of America. We analyzed sentiment score over time, and summarize the frequency of Positive and Negative terms.



sentiment over time

### Method used

We used 3 methods to conduct Sentiment Analysis: a) look up dictionary; b) Machine learning; c) Combine the above two.  
We can see the difference by plotting them in one graph:

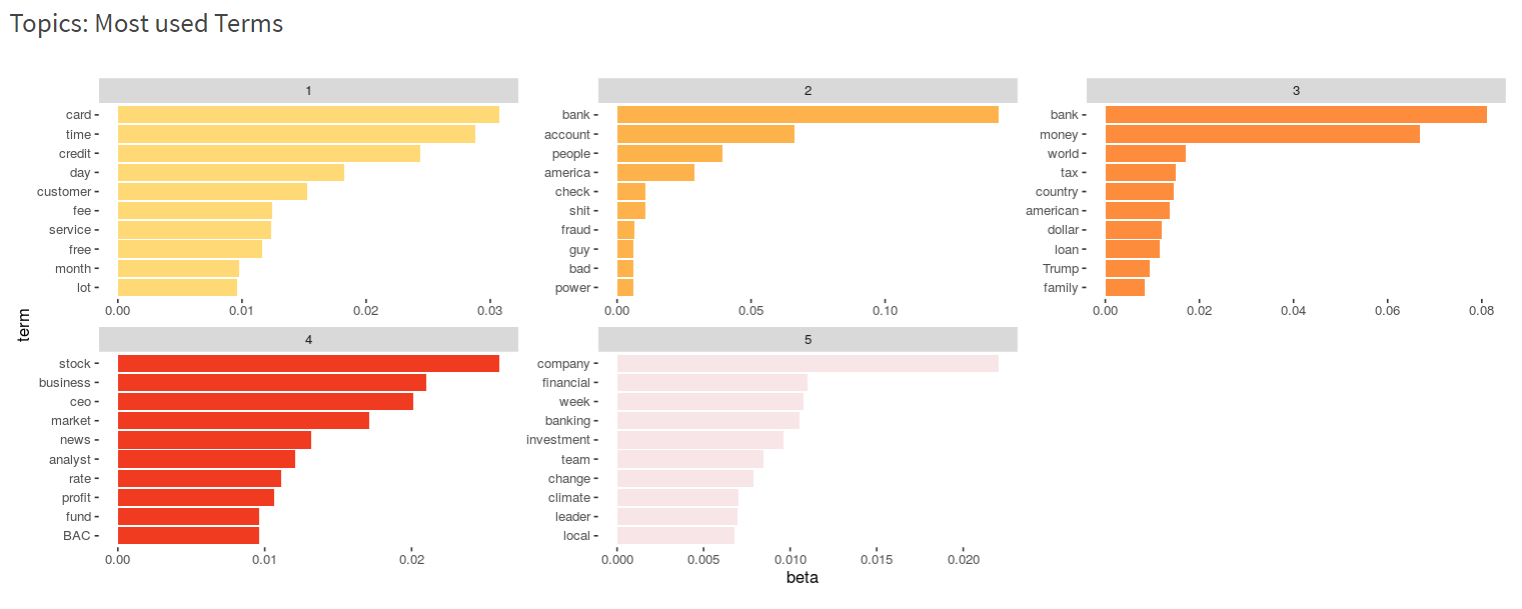


Sentiment on different methods

## Topic Modelling

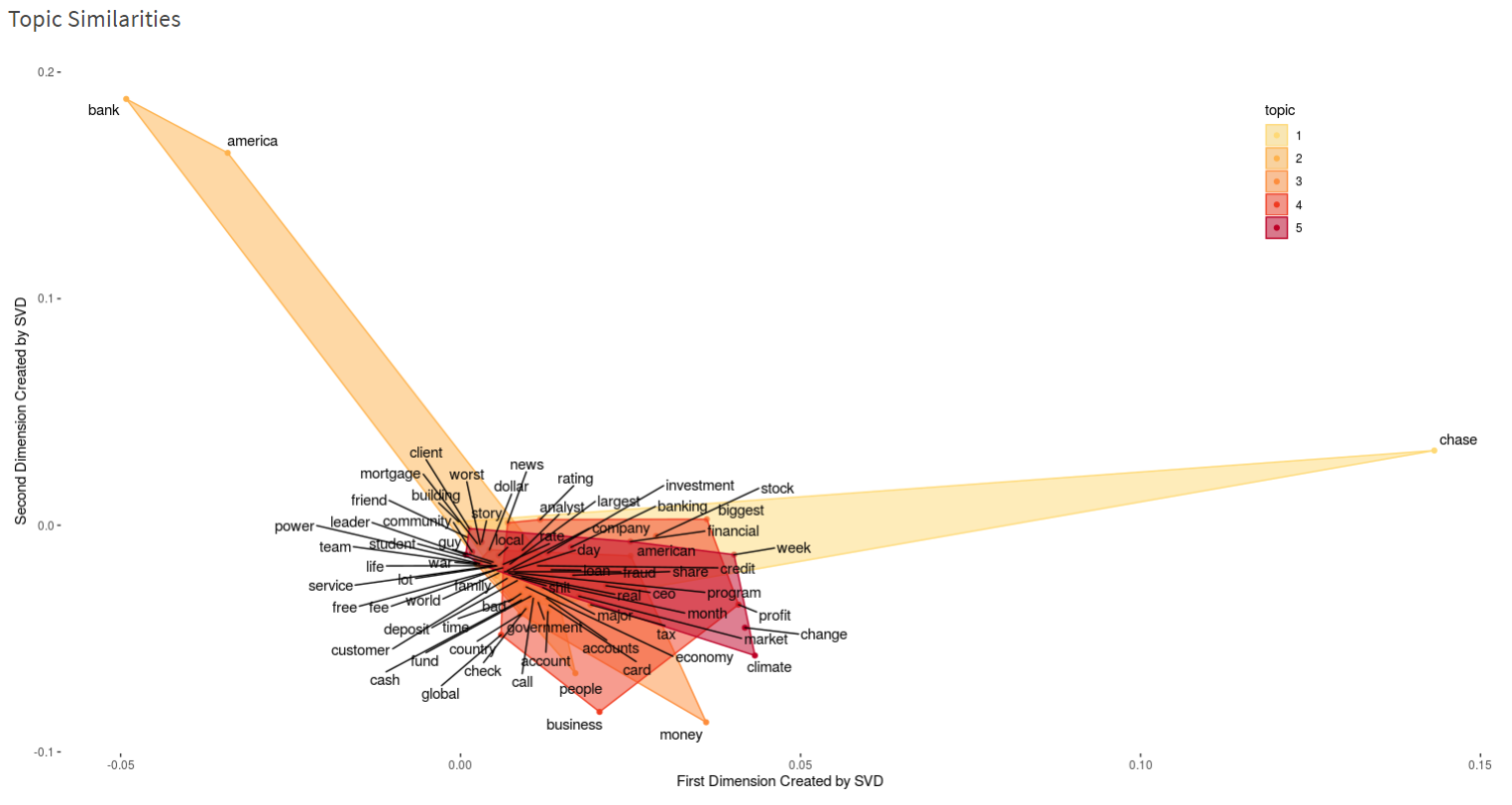
To have an insight into what topic do people talk about when they mention Bank of America, we conducted topic modelling on tweets people mention Bank of America.

We have identified 5 relevant topics among the tweets and frequency of terms in each topic.



topics and terms.

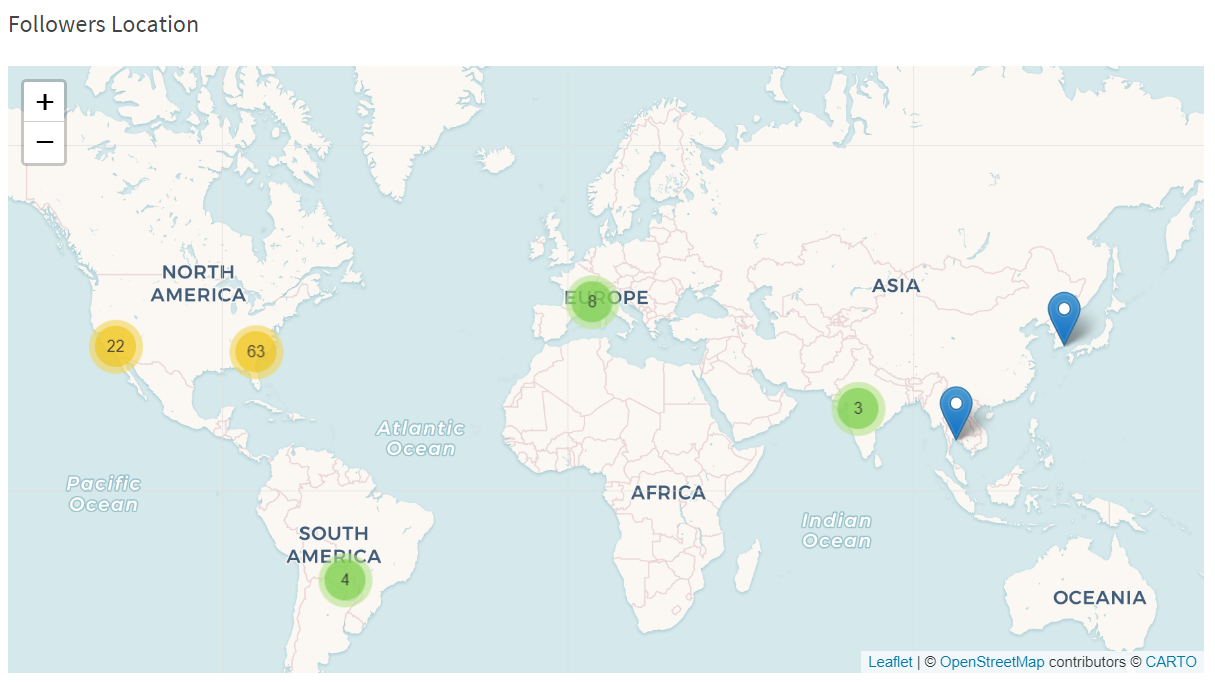
We can also compare the similarity of the top terms of the topics. The words are displayed in a 2-D plane (vectorized) with a SVD model. Words that are similar according to the SVD are displayed in closer distance to each other.



topics similarity.

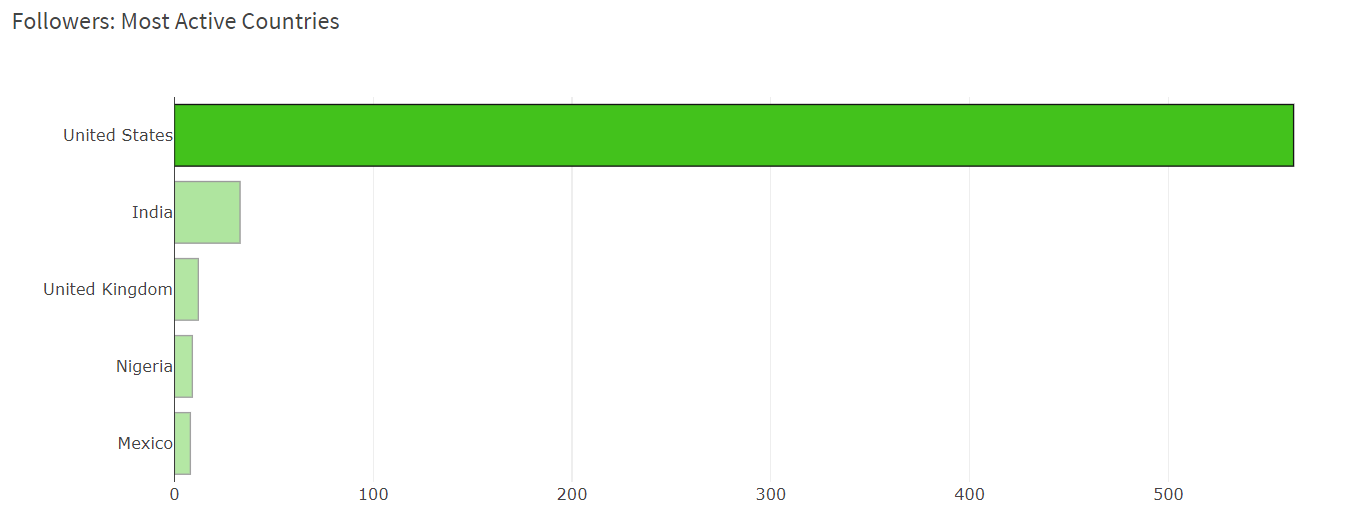
## Follower Profile

To get to know the people following Bank of America, we collected followers’ profile and summarized their twitter behavior.  
We can see where are the followers from. *Only the followers who shared their location on their profile*



location map.

We can also see which country’s followers are highly active on Twitter.



ranking of country.